

Esercitazione excel avanzato

Introduzione

In questa esercitazione imparerete a importare dati da un file di testo in un foglio di calcolo e ad utilizzare le funzioni di excel utili per l'elaborazione di grandi quantità di informazioni:

- filtri semplici
- filtri complessi
- calcolo di subtotali
- tabelle pivot

Gli esercizi proposti si basano su di una tabella di oltre 2500 righe, i cui valori sono memorizzati in un file di testo chiamato "dati_da_importare.txt".

La tabella contiene informazioni relative agli iscritti ad un corso di laurea negli anni compresi tra il 1988 e il 1995.

La tabella è composta dalle seguenti colonne:

- anno (anno accademico)
- sesso (0 se di sesso maschile, 1 se di sesso femminile)
- eta'
- peso (espresso in Kg)
- altezza (espressa in cm)
- dipsci (1 se lo studente ha conseguito la maturità scientifica, 0 altrimenti)
- dipcla (1 se lo studente ha conseguito la maturità classica, 0 altrimenti)
- diptec (1 se lo studente è diplomato presso un istituto tecnico, 0 altrimenti)
- dipalt (1 se lo studente ha un altro tipo di diploma)
- compon (numero di componenti del nucleo familiare)
- occhiali (1 se lo studente porta gli occhiali, 0 altrimenti)
- fumo (1 se lo studente è fumatore, 0 altrimenti)

Importazione dati

Scopo: importare i dati dal file dati_da_importare.txt

Nelle esercitazioni didattiche spesso vengono utilizzate tabelle di piccole dimensioni, contenenti solo poche righe e i cui dati vengono immessi a mano. Nell'utilizzo reale, i fogli di calcolo vengono utilizzati per trattare grandi volumi di dati, che sono importati da sorgenti esterne (file, basi di dati, ...). Sorgenti esterne diverse possono avere formati diversi, che non sono supportati nativamente da Excel. Tuttavia la grandissima maggioranza delle applicazioni consentono di esportare ed importare i dati in un formato standard, chiamato CSV. CSV significa Comma Separated Values (valori separati da virgole), e consiste in un file di testo semplice (solo caratteri ASCII, senza formattazioni particolari) contenente valori (sia testo che numeri) separati da virgole.

In questo esercizio creerete una tabella di medie dimensioni a partire dai dati

memorizzati in formato CSV nel file "dati_da_importare.txt".

- aprire un nuovo foglio di calcolo (vuoto)
- selezionare Data -> Import external data -> Import data
- selezionare il file dati_da_importare.txt (l'esatta posizione del file nel vostro filesystem dipende a seconda di dove avete scompattato l'archivio)
- scegliere "delimited" e importare tutti i dati contenuti nel file dati_da_importare.txt (compresa la prima riga), premere next
- scegliere la virgola come delimitatore
- selezionare il foglio e la cella in cui importare, premere ok
- risultato: nel foglio di lavoro corrente compare la tabella con tutti i dati che saranno utilizzati per gli esercizi successivi. La tabella ottenuta deve avere le 12 colonne descritte nell'introduzione e 2760 righe (compresa la riga dell'intestazione)

Filtri semplici

Scopo: selezionare solo gli studenti iscritti nel 1988 di sesso maschile e dal peso di 62 kg

La tabella ottenuta dall'esercizio precedente contiene molti dati, difficili da analizzare. In particolare è complicato cercare all'interno della tabella solo i pochi campi che corrispondono a ben determinati criteri di ricerca.

In questo esercizio applicherete dei filtri semplici alla tabella, in modo da poter effettuare facilmente delle semplici ricerche.

- selezionare una cella all'interno della tabella dei dati
- selezionare Data -> Filter -> AutoFilter
- utilizzare i menu a tendina che compaiono nella prima riga della tabella per filtrare i dati in base ai vari campi
 - nel menu a tendina della colonna anno selezionare 1988
 - nel menu a tendina della colonna sesso selezionare 0
 - nel menu a tendina della colonna peso selezionare 62
- risultato: di tutte le righe contenute nella tabella vengono visualizzate solo quelle che soddisfano tutti e tre i vincoli selezionati nel punto precedente, corrispondenti alla seguente espressione in logica booleana:

((ANNO = 1988) AND (SESSO = 0) AND (PESO = 62))

La tabella risultante è la seguente:

| ANNO | SESSO | ETA | PESO | ALTEZZA | DIPSCI | DIPCLA | DIPTEC | DIPALT | COMPON | OCCHIALI | FUMO |
|------|-------|------|------|---------|--------|--------|--------|--------|--------|----------|------|
| 1988 | 0 | 20.8 | 62 | 180 | 0 | 0 | 1 | 0 | 5 | 0 | 1 |
| 1988 | 0 | 20.6 | 62 | 176 | 0 | 0 | 1 | 0 | 5 | 0 | 0 |

- verificare gli effetti di altre regole di filtro cambiando i valori nei menu a tendina delle colonne della tabella

- eliminare tutte le regole di filtro impostate prima di procedere al prossimo esercizio selezionando Data -> Filter -> Show All

Filtri complessi

Scopo: selezionare tutti i maschi di altezza inferiore a 160 cm e tutte le femmine di altezza superiore a 177 cm

Funzione logica booleana:

(((sesso = 0) AND (altezza < 160)) OR ((sesso=1) AND (altezza > 180)))

Mediante i menu a tendina messi a disposizione dai filtri semplici, non è possibile esprimere funzioni logiche di filtro come quella descritta nello scopo di questo esercizio. Per risolvere questo esercizio è necessario utilizzare i filtri complessi.

In un filtro complesso, le regole di filtro vengono lette da una tabella, in cui vengono inseriti tutti i criteri di filtraggio. Tutte le condizioni inserite sulla stessa riga della tabella vengono considerate in AND logico, mentre le condizioni rappresentate da righe diverse vengono considerate in OR.

La funzione booleana di questo esercizio può essere scritta come

(riga_1 OR riga_2)

dove a riga_1 dobbiamo sostituire

((sesso = 0) AND (altezza < 160))

e a riga_2 dobbiamo sostituire

((sesso = 1) AND (altezza > 180))

- costruite SOPRA LA TABELLA CON I DATI una tabellina con i criteri di filtraggio (se non avete abbastanza spazio sopra la tabella, createlo inserendo delle righe vuote)

| SESSO | ALTEZZA |
|-------|---------|
| 0 | < 160 |
| 1 | > 180 |

- potete verificare che questa tabella descrive le regole di filtro desiderate (la prima riga esprime le condizioni di riga_1, la seconda riga esprime le condizioni di riga_2, le due righe sono in OR logico tra loro)
- selezionare la prima cella della tabella dei dati
- selezionare Data -> Filter -> Advanced Filter
- comparire la mascherina che richiede l'inserimento dei dati necessari alla creazione del filtro
- controllare che il campo List Range contenga l'area della tabella con i dati
- nel campo Criteria Range inserire l'area della tabellina contenente i criteri di filtraggio
- premere OK

- risultato: di tutte le righe contenute nella tabella vengono visualizzate solo quelle che soddisfano l'espressione booleana che è rappresentata dalla tabellina dei criteri di filtraggio.

La tabella risultante è la seguente:

| ANNO | SESSO | ETA | PESO | ALTEZZA | DIPSCI | DIPCLA | DIPTEC | DIPALT | COMPON | OCCHIALI | FUMO |
|------|-------|------|------|---------|--------|--------|--------|--------|--------|----------|------|
| 1989 | 0 | 41.1 | 61 | 155 | 0 | 0 | 1 | 0 | 5 | 1 | 0 |
| 1992 | 0 | 19.3 | 50 | 158 | 1 | 0 | 0 | 0 | 4 | 0 | 0 |
| 1994 | 1 | 25.1 | 73 | 182 | 0 | 1 | 0 | 0 | 4 | 0 | 1 |

- eliminare tutte le regole di filtro impostate prima di procedere al prossimo esercizio selezionando Data -> Filter -> Show All

Subtotali

Scopo: creare la seguente tabella in un nuovo foglio di calcolo chiamato "altezze medie", e calcolare i dati mancanti (contrassegnati dal punto interrogativo)

| anno | altezza media | altezza media maschi | altezza media femmine |
|------|---------------|----------------------|-----------------------|
| 1988 | ? | ? | ? |
| 1991 | ? | ? | ? |
| 1995 | ? | ? | ? |

Per applicare delle funzioni (in questo esempio, la funzione MEDIA()) solo ad alcune righe di una tabella, occorre utilizzare la funzione SUBTOTALE(). Questa funzione prende due parametri: un numero intero e un intervallo di celle. Il numero intero rappresenta la funzione da utilizzare (9 indica SOMMA, 3 indica CONTAVALORI, 1 indica MEDIA, ...). La lista di tutte le funzioni disponibili è consultabile utilizzando la guida di Excel. Il secondo parametro rappresenta l'intervallo di celle a cui applicare la funzione.

La caratteristica interessante di SUBTOTALE() è che la funzione indicata dal primo parametro viene applicata esclusivamente alle celle che sono visibili, mentre tutte le celle che sono nascoste (ad esempio per la non corrispondenza con le espressioni di filtro impostate) vengono ignorate.

Questo esercizio può essere risolto utilizzando la funzione SUBTOTALE() e impostando delle regole di filtro per visualizzare solo le celle che interessano (ad esempio: l'altezza media dei maschi nel 1995 può essere calcolata con SUBTOTALE() applicando la funzione 3 (MEDIA) su tutte le celle della colonna delle altezze e utilizzando dei filtri semplici per selezionare esclusivamente i maschi (SESSO=0) del 1995).

- selezionare Insert -> Worksheet
- rinominare il nuovo foglio di lavoro in "altezze medie"
- creare lo "scheletro" della tabella (scrivere manualmente la prima riga e la prima colonna)
- tornare al foglio di calcolo "dati_da_importare"
- selezionare una cella SOPRA LA TABELLA CON I DATI
- scrivere la formula =SUBTOTAL(1,F11:F2769) [sostituire al secondo parametro l'intervallo di celle contenente i valori delle altezze della vostra tabella dei dati]
- selezionare Data -> Filter -> AutoFilter
- il risultato della formula è calcolato solo in base alle celle visualizzate
 - es: per calcolare l'altezza media delle femmine nel 1988 selezionare solo le righe con anno 1988 e sesso 1
- agendo sui filtri calcolare tutti i valori necessari e riportare i dati calcolati nella tabella del foglio "altezze medie"
- risultato:

| anno | altezza media | altezza media maschi | altezza media femmine |
|------|---------------|----------------------|-----------------------|
| 1988 | 168.739 | 175.759 | 163.868 |
| 1991 | 169 | 176.144 | 164.185 |
| 1995 | 170.301 | 177.343 | 164.656 |

Scopo: creare la seguente tabella in un foglio di calcolo chiamato "numero iscritti"

| anno | numero iscritti | numero maschi | numero femmine |
|------|-----------------|---------------|----------------|
| 1988 | ? | ? | ? |
| 1991 | ? | ? | ? |
| 1995 | ? | ? | ? |

- procedere in modo analogo all'esercizio precedente
- indizio: nella funzione subtotal(), usare la funzione contavalori (3) al posto della funzione somma (9)
- risultato:

| anno | numero iscritti | numero maschi | numero femmine |
|------|-----------------|---------------|----------------|
| 1988 | 476 | 281 | 195 |
| 1991 | 327 | 132 | 195 |
| 1995 | 236 | 105 | 131 |

Tabelle pivot

Scopo: creare in un nuovo foglio di calcolo una tabella pivot che riassume la tabella con i dati e che permetta di visualizzare facilmente quanti studenti erano in possesso della maturità classica e scientifica suddividendoli per sesso e per anno di iscrizione

Le tabelle pivot sono uno strumento estremamente utile per aggregare grandi moli di dati, e consentono di avere una visione complessiva di informazioni che sarebbero difficilmente interpretabili.

In generale, una tabella pivot comprende tre parti principali: l'area delle righe, l'area delle colonne e l'area dei dati. Nell'area delle righe e delle colonne possono essere messi i campi che si vogliono utilizzare suddividere i dati, mentre nell'area dei dati devono essere inseriti i campi che si vogliono aggregare.

- selezionare la prima cella della tabella con i dati
- selezionare Data -> Pivot Table and Pivot Chart Reports
- selezionare le voci "Microsoft Office Excel list or database" e "Pivot table", premere Next
- controllare che il campo Range contenga l'area della tabella con i dati, premere Next
- selezionare Layout (tasto in basso a sinistra)
- trascinare ANNO nell'area delle righe
- trascinare SESSO nell'area delle colonne
- trascinare DIPSCI e DIPCLA nell'area dei dati (DATA)
- premere Next
- selezionare la voce "New Worksheet" e premere Finish
- risultato: una tabella pivot simile a quella riportata nella seguente immagine

| | | SESSO | | |
|---------------------|---------------|-------|-----|-------------|
| ANNO | Data | 0 | 1 | Grand Total |
| 1988 | Sum of DIPSCI | 54 | 57 | 111 |
| | Sum of DIPCLA | 50 | 111 | 161 |
| 1989 | Sum of DIPSCI | 27 | 36 | 63 |
| | Sum of DIPCLA | 37 | 75 | 112 |
| 1990 | Sum of DIPSCI | 45 | 38 | 83 |
| | Sum of DIPCLA | 24 | 81 | 105 |
| 1991 | Sum of DIPSCI | 37 | 40 | 77 |
| | Sum of DIPCLA | 20 | 69 | 89 |
| 1992 | Sum of DIPSCI | 35 | 45 | 80 |
| | Sum of DIPCLA | 25 | 88 | 113 |
| 1993 | Sum of DIPSCI | 31 | 37 | 68 |
| | Sum of DIPCLA | 27 | 77 | 104 |
| 1994 | Sum of DIPSCI | 47 | 41 | 88 |
| | Sum of DIPCLA | 18 | 69 | 87 |
| 1995 | Sum of DIPSCI | 32 | 35 | 67 |
| | Sum of DIPCLA | 12 | 38 | 50 |
| Total Sum of DIPSCI | | 308 | 329 | 637 |
| Total Sum of DIPCLA | | 213 | 608 | 821 |

- potete modificare l'aspetto della tabella agendo sui menu a tendina e potete modificare la struttura della tabella pivot utilizzando la finestra Pivot Table Field List
 - es: aggiungere il dato DIPTEC all'area dei dati della tabella pivot per visualizzare anche i dati relativi agli studenti diplomati all'istituto tecnico